

DATABANK MERGE UPDATE (20120214)

The following are some tables and graphics generated as a result of two different versions of the databank merge program. Currently the program focuses on Maximum Temperature, and each version of the merge applies a different algorithm when making comparisons of data with 60 months (or greater) of data overlap:

- a) NRMSD (Normalized Root Mean Squared Deviation)
- b) IA (Index of Agreement)

Part a) NRMSD

$$NRMSD = \frac{\sqrt{\frac{\sum_{i=1}^n (M_i - C_i)^2}{n}}}{MAX - MIN}$$

M_i = master temperature data for a specific year/month

C_i = candidate temperature data for a specific year/month

n = number of common months between master/candidate

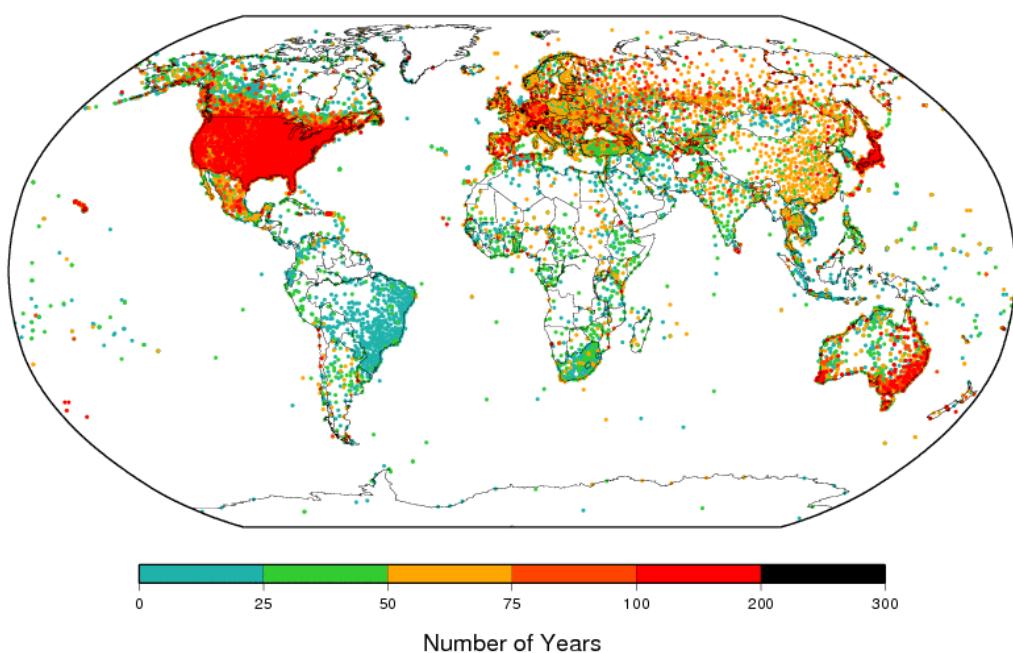
MAX = maximum temperature data out of BOTH M and C

MIN = minimum temperature data out of BOTH M and C

- 1) Number of Stations, color coded by period of record

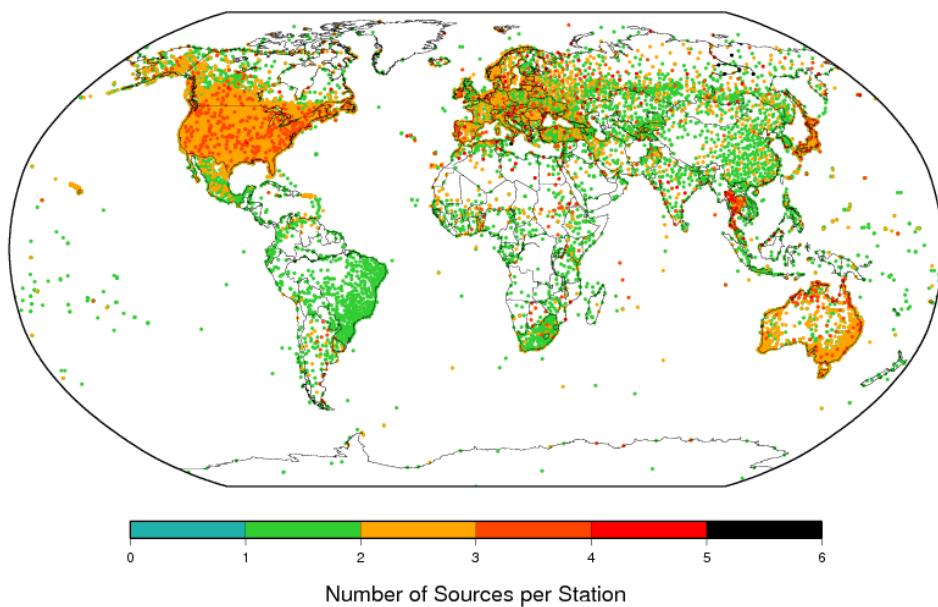
Stage3 TMAX (Date=20120214, Method=NRMSD) UNOFFICIAL

Number of UNIQUE Station Records: 29639

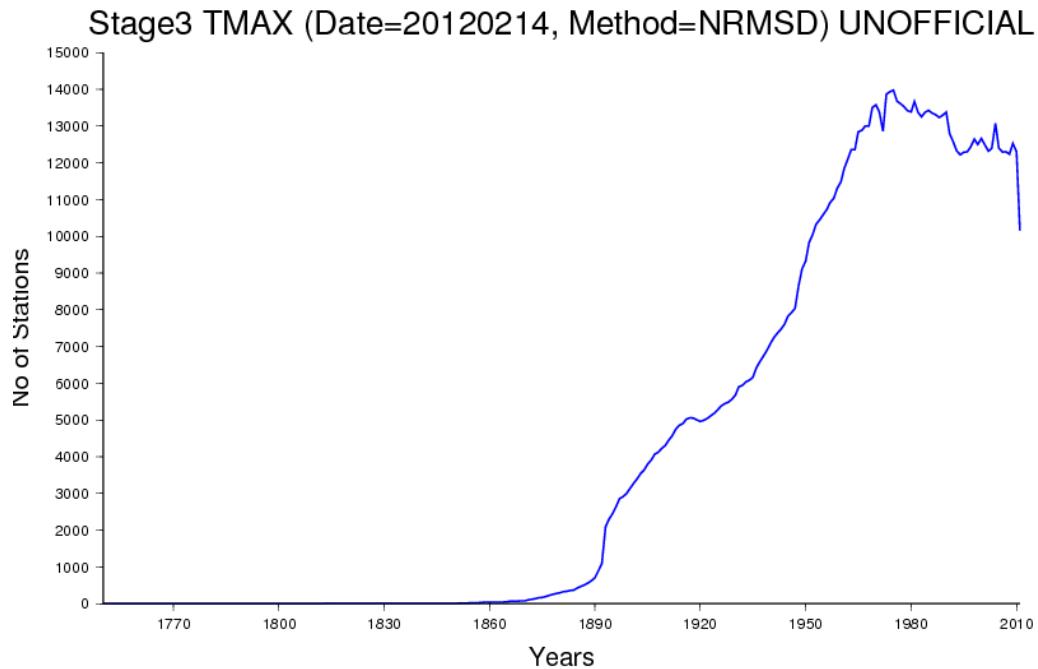


- 2) Number of Stations, color coded by number of sources per station

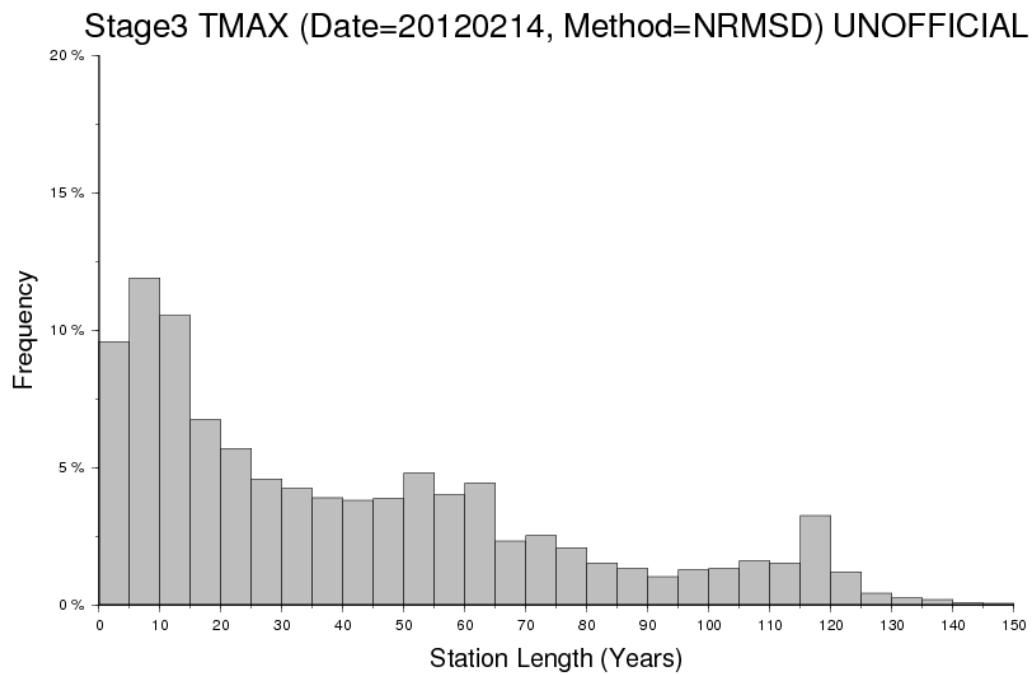
Stage3 TMAX (Date=20120214, Method=NRMSD) UNOFFICIAL
Number of UNIQUE Station Records: 29639



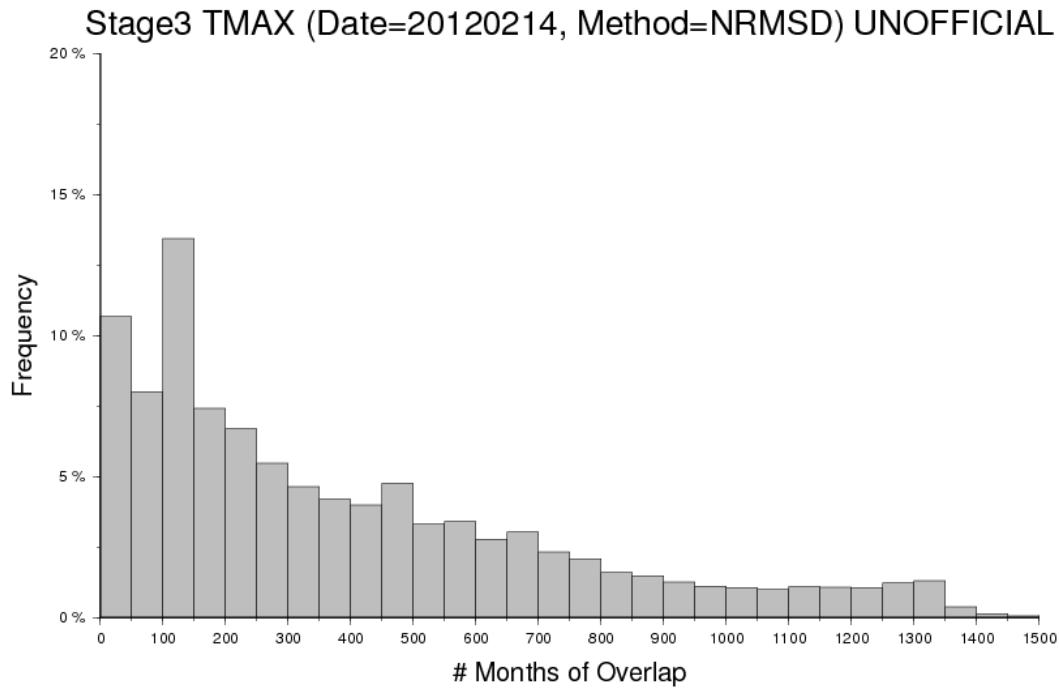
- 3) Time series of number of stations in Stage 3 Databank



4) Histogram of station length (in years)



5) Histogram of overlap period



6) Distribution of results using NRMSD

DATASET NAME	# STATIONS	% MERGED	% UNIQUE	% THRASHED
mexico	95	0	95.79	4.21
vietnam	32	0	78.12	21.88
usforts	217	6.91	4.61	88.48
channel-islands	2	0	50	50
ecuador	1	0	100	0
pitcairnisland	3	0	33.33	66.67
beyrout	1	0	100	0
brazil	495	0	97.78	2.22
wwr	5056	18.26	35.56	46.18
colonialera	1020	4.8	63.73	31.47
east-africa	263	7.22	34.6	58.17
antarctica-southpole	1	0	100	0
ispd-swiss	3	0	0	100
ispd-ipy	2	0	50	50
ispd-sydney	1	0	0	100
antarctica-scar-reader	44	0	56.82	43.18
spain	22	27.27	0	72.73
russia	517	28.05	30.95	41.01
uruguay-inia	5	0	20	80
swiss-digihom	3	0	0	100
ispd-tunisia-morocco	13	7.69	61.54	30.77
ecaknmi	1350	30.59	0.52	68.89
sacaknmi	172	63.37	1.16	35.47
japan	157	0	100	0
ukmet-hist	37	24.32	56.76	18.92
knmi	1551	50.87	24.82	24.31
russsource	36158	67.7	7.53	24.77
ghcnsource	15585	0	17.89	82.11
ghcnmv2	13499	30.09	4.65	65.26
central-asia	234	0	50	50
canada	337	79.23	0	20.77
australia	103	70.87	0	29.13
arctic	133	0	30.83	69.17
histalp	138	0	22.46	77.54
hadcru3	5113	0	14.92	85.08

Part b) Index of Agreement

M_i = master temperature data for a specific year/month

C_i = candidate temperature for a specific year/month

\bar{M} = Mean of entire master temperature

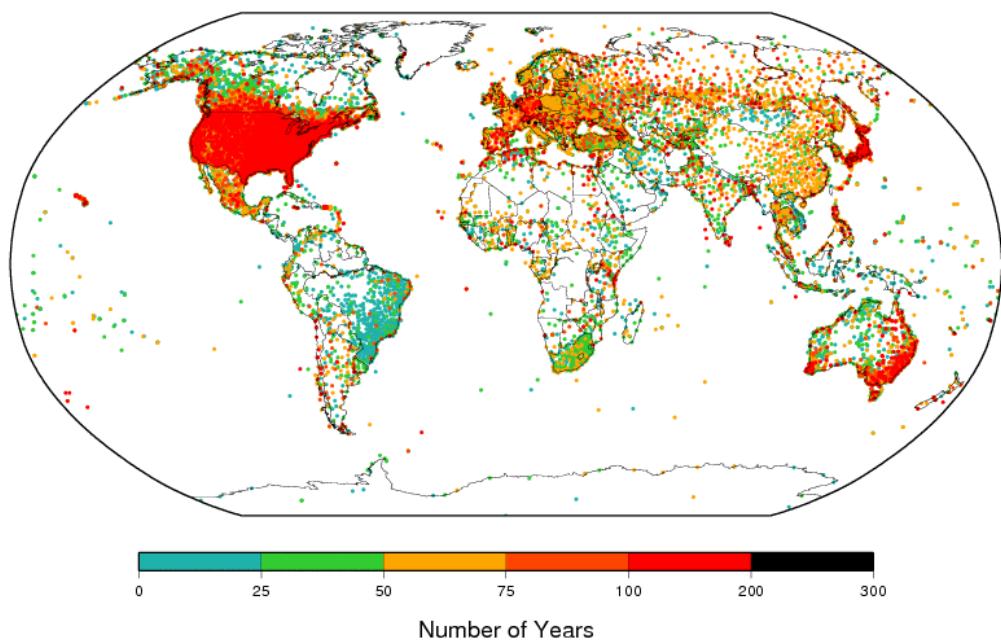
n = number of common months between master/candidate

$$IA = 1.0 - \frac{\sum_{i=1}^n |M_i - C_i|}{\sum_{i=1}^n (|C_i - \bar{M}| + |M_i - \bar{M}|)}$$

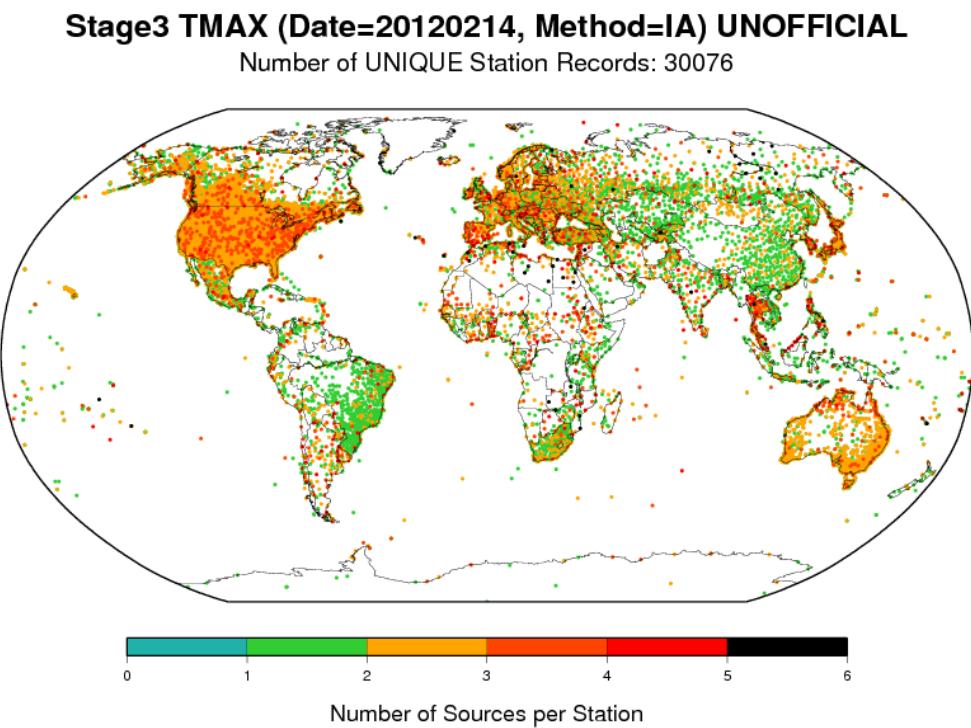
- 1) Number of Stations, color coded by period of record

Stage3 TMAX (Date=20120214, Method=IA) UNOFFICIAL

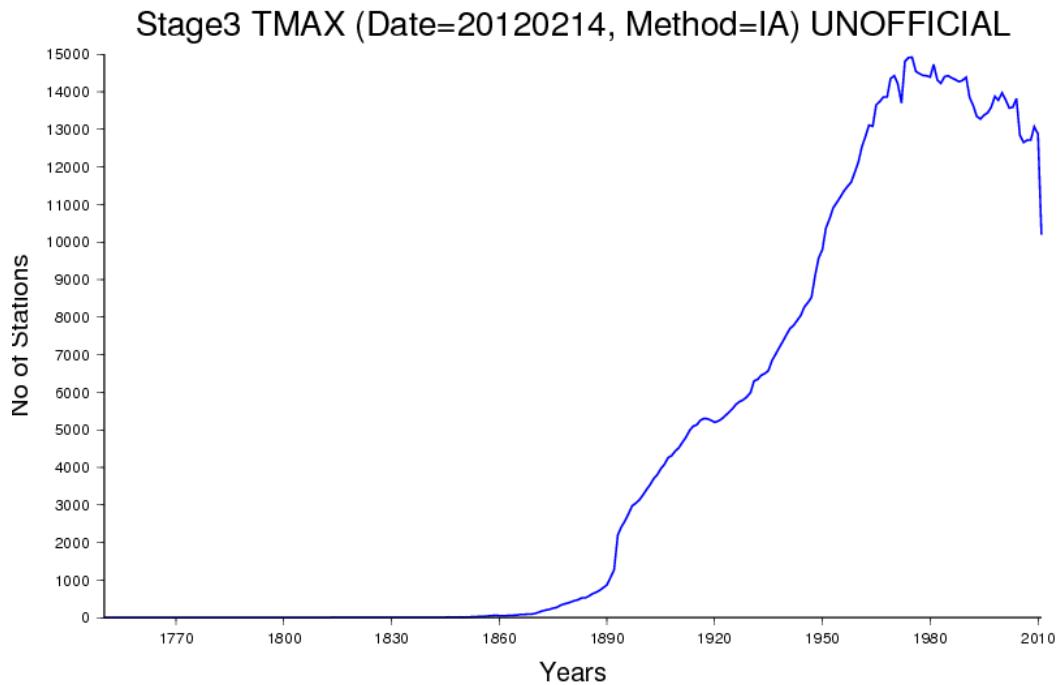
Number of UNIQUE Station Records: 30076



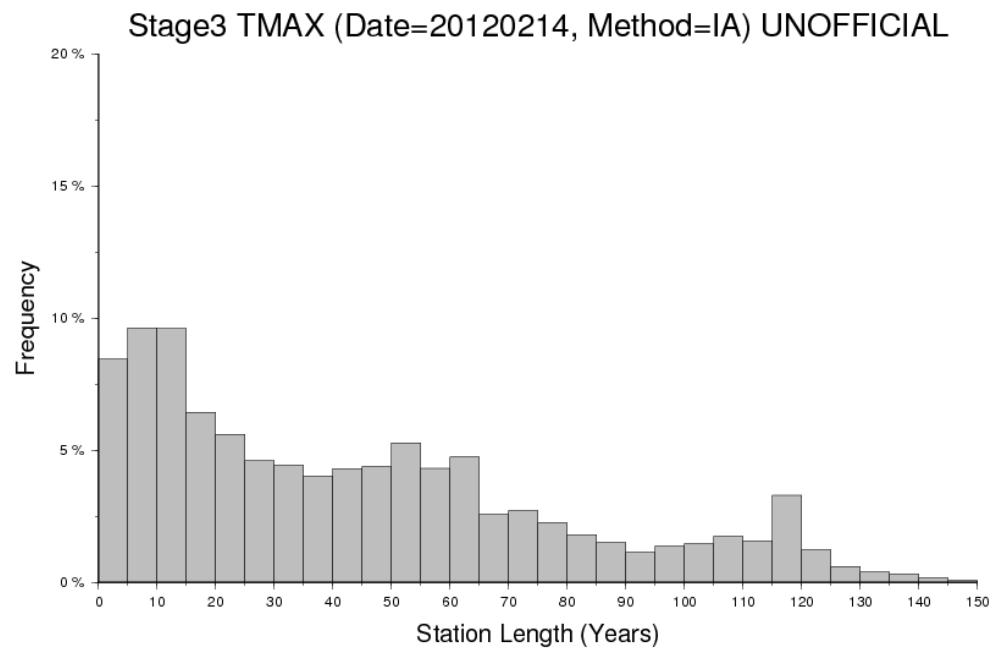
- 2) Number of Stations, color coded by number of sources per station



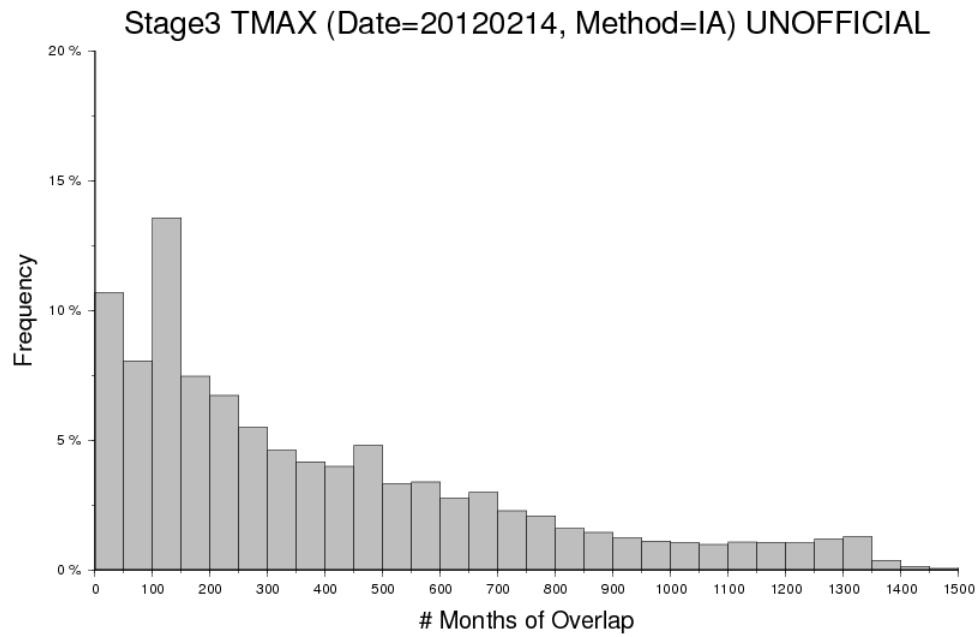
- 3) Time series of number of stations in Stage 3 Databank



4) Histogram of Station length (in years)



5) Histogram of overlap period



6) Distribution of results using Index of Agreement

DATASET NAME	# STATIONS	% MERGED	% UNIQUE	% TRASH
mexico	95	3.16	95.79	1.05
vietnam	32	0	78.12	21.88
usforts	217	94.93	4.61	0.46
channel-islands	2	50	50	0
ecuador	1	0	100	0
pitcairnisland	3	0	33.33	66.67
beyrout	1	0	100	0
brazil	495	0	97.78	2.22
wwr	5056	59.41	35.56	5.02
colonialera	1020	29.71	63.82	6.47
east-africa	263	51.71	38.4	9.89
antarctica-southpole	1	0	100	0
ispd-swiss	3	100	0	0
ispd-ipy	2	50	50	0
ispd-sydney	1	100	0	0
antarctica-scar-reader	44	43.18	56.82	0
spain	22	77.27	0	22.73
russia	517	41.01	30.95	28.05
uruguay-inia	5	60	20	20
swiss-dighom	3	100	0	0
ispd-tunisia-morocco	13	38.46	61.54	0
ecaknmi	1350	55.19	0.52	44.3
sacaknmi	172	66.28	1.74	31.98
japan	157	0	100	0
ukmet-hist	37	40.54	56.76	2.7
knmi	1551	69.76	24.95	5.29
russsource	36158	88.55	7.8	3.64
ghcnsource	15585	80.57	17.88	1.55
ghcnmv2	13499	92.09	4.68	3.23
central-asia	234	43.59	55.98	0.43
canada	337	85.76	0	14.24
australia	103	64.08	0	35.92
arctic	133	68.42	30.83	0.75
histalp	138	77.54	22.46	0
hadcru3	5113	83.04	14.9	2.05